Dominance solvability of dynamic bargaining games

TYSON, CJ

Dominance Solvability of Dynamic Bargaining Games

Christopher J. Tyson

Department of Economics
Queen Mary, University of London

January 12, 2009

Abstract

We formulate and study a general finite-horizon bargaining game with simultaneous moves and a disagreement outcome that need not be the worst possible result for the agents. Conditions are identified under which the game is dominance solvable in the sense that iterative deletion of weakly dominated strategies selects a unique outcome. Our analysis uses a backward induction procedure to pinpoint the latest moment at which a coalition can be found with both an incentive and the authority to force one of the available alternatives. Iterative dominance then implies that the alternative characterized in this way will be agreed upon at the outset — or, if a suitable coalition is never found, that no agreement will be reached.

JEL classification codes: C78, D71, D74.

Keywords: backward induction, coalition, core, weak dominance.

1 Introduction

1.1 Motivation and objectives

Since its early application to dynamic bargaining in the work of Stahl [24] and Rubinstein [20], game-theoretic modeling has led to a better understanding of the effects on negotiated outcomes of voting rules, outside options, private information, and other such factors. At the same time, however, a number of difficulties have arisen that cast doubt upon the robustness and predictive power of results in this area.

Two such difficulties are relevant here. Firstly, the bargaining protocols used in most theoretical and applied studies have at their heart an assumption of "temporal monopoly [power]" that has never been adequately defended.[1] And secondly, quite plausible models of negotiation can turn out to possess large sets of equilibria that are not always readily

[1] For discussion and critiques of this assumption, see Kreps [14, pp. 563–565], Smith and Stacchetti [23], and Simsek and Yildiz [22].

pruned by any compelling selection criterion. For conciseness let us refer to these as the "monopoly" and "multiplicity" problems, respectively.

In this paper we shall analyze a bargaining model that deals, *to a limited extent*, with both of the problems just mentioned. To combat the monopoly problem, we shall posit a simultaneous voting protocol that is symmetric among the participants and thus avoids allocating bargaining power arbitrarily in the form of permission to commit or to delay committing to an action. When similar "simultaneous-offer" games have been studied in the past (e.g., by Nash [18] and Chatterjee and Samuelson [6]), the conclusion (e.g., of Dekel [8, p. 301]) has been that they quickly run afoul of the multiplicity problem. But this can be handled, we shall find, by using dominance solvability rather than any variety of strategic equilibrium as our solution concept.

The combination of simultaneous moves and dominance analysis will allow us to work with a model that is in several respects exceedingly general. Our setting will encompass bargaining environments with any (finite) number of agents, an arbitrary (finite) set of alternatives, time preferences subject only to weak regularity assumptions, and a variety of different rules for reaching an agreement. The treatment of multilateral environments in particular is noteworthy because these have been shown (e.g., by Sutton [25, pp. 721–723] and Baron and Ferejohn [3, pp. 1189–1190]) to be highly susceptible to the multiplicity problem.[2] We shall, however, employ one structural assumption that is clearly restrictive: The negotiation must have a finite horizon so that backward induction can be used to organize the deletion of dominated strategies.

Our main result (Theorem 3.16) gives sufficient conditions for dominance solvability of the bargaining game in question and, when these hold, identifies the implied outcome of the interaction. *No claim is made that the conditions are weak or likely to be satisfied in most situations of interest.*[3] Indeed, it will be easy to exhibit economically-relevant settings in which they fail. Our objectives, therefore, are modest ones: to show that the bargaining situations in a circumscribed but not negligible class are dominance solvable, and to use our somewhat nonstandard model to gain a new perspective on the monopoly and multiplicity problems mentioned above.

1.2 An illustrative example

A trade liberalization conference has been called for a regional organization consisting of the nations A, B, C, D, and E.[4] Representatives of these countries will debate which of two draft treaties, a or b, should be sent to their respective legislatures for ratification. The negotiations will begin at nine o'clock on Monday morning and will end either when an agreement is reached or (in the absence of an earlier agreement) at five o'clock on Friday afternoon. There will be eight hours of discussions per day, and hence (up to) forty hours

[2] This being said, other escape routes from the multiplicity problem have been found in the multilateral case. For example, Banks and Duggan [2] and Cho and Duggan [7] obtain core equivalence results in the context of the Baron-Ferejohn model by endowing the alternative set with a one-dimensional structure.

[3] We shall therefore distinguish these "conditions" from our less objectionable "assumptions" regarding various regularity or technical (e.g., genericity) properties of the model. Note that any assumption is a tacit hypothesis of each result that follows it in the text, whereas conditions are appealed to explicitly.

[4] For later reference, note that this illustration is formalized in Example 4.2 below.

	A	B	C	D	E
draft a	64	48	32	16	-16
draft b	-8	8	48	56	64
no agreement	40	40	40	40	40

Table 1: Preferences in the trade liberalization problem. Displayed are the payoffs (before bargaining costs) to the five agents from the two draft treaties and from disagreement.

in total over the course of the week. According to the bylaws of the organization, each country's opinion is to be given equal weight and a supermajority of four votes is needed to reach an agreement.

The payoffs to the five agents from the three possible substantive outcomes are shown in Table 1. For example, agent C will receive 32 if a is agreed, 48 if b is agreed, and 40 if there is no agreement. In addition, each agent will lose one unit of payoff per hour of negotiations, reflecting the opportunity cost of being in the conference center rather than on the nearby beach. Note that — crucially — the default ("no agreement") outcome is not available to the bargainers until the end of the day on Friday, as agreeing to disagree and heading to the beach before the end of the conference would be viewed as a dereliction of duty by their superiors. And note also that the net payoff to each agent in the event of default is $40 - 40 = 0$, a convenient normalization.

How should we expect this scenario to play out? Observe first of all that there are two countries (A and B) that prefer draft a to draft b and three (C, D, and E) with the opposite preference. Thus, in view of the need for four votes, the problem is a nontrivial one. Can we conclude then that no agreement will be reached, or does one or the other faction possess sufficient bargaining power to implement its preferred alternative?

Since the negotiations have a finite horizon, game theory offers the tool of backward induction: The agents' behavior in earlier contingencies will depend on what they expect to happen in later ones, so we should think first about the end of the week and proceed backwards.

1. At any time after thirty-two cumulative hours of meetings (i.e., any time on Friday), we can be certain that no agreement will be reached. After this point each agent's payoff from his dispreferred alternative will be less than the disagreement payoff, and thus no agent will be tempted to switch sides in order to gain release from the conference room.

2. Between sixteen and thirty-two hours, agents A, B, D, and E will remain unwilling to switch sides. Agent C, however, *will* be tempted to switch since alternative b is clearly unattainable and a is preferable to default (which by Step 1 is the foreseeable consequence of a failure to agree by the end of thirty-two hours). But since A, B, and C together lack the power to force a, we should again expect disagreement.

3. Just before the end of sixteen hours, agents A, B, and E will continue to hold firm while D will join C in being willing to defect. And since A, B, C, and D together make up the requisite supermajority, we should expect agreement on alternative a late Tuesday afternoon.

4. Marginally earlier on Tuesday, the "continuation outcome" that can be anticipated absent an immediate agreement is (by Step 3) that a will be agreed upon a moment later. Agreement on a earlier rather than later is better for everyone, and there is no prospect of choosing b once it is understood by A and B that a can be obtained after a short delay. Thus a will be agreed upon at the earlier moment and hence — applying the same logic inductively — at all moments before the end of sixteen hours. In particular, we should expect a to be agreed upon immediately.

Observe that this analysis of the trade liberalization problem hinges upon a critical moment (namely, the end of the sixteenth hour of negotiations) after which disagreement is certain and just before which it is possible to assemble a coalition (consisting of A, B, C, and D) whose members collectively have the power to implement one of the alternatives (namely, a) and individually have an incentive to do so rather than waiting for the default outcome. To identify these features of the problem, we can define the value of an alternative to an agent as the latest time at which he prefers it to eventual default (e.g., C values a at thirty-two hours). The strength of a given coalition is then the smallest of its members' valuations, and the coalition that is strongest in this sense "wins" by having its associated alternative selected immediately.

Of course, not all bargaining problems will yield so easily to this method. For example, if each entry in the first two rows of Table 1 were to exceed 40 — making disagreement the worst possible outcome for all agents — then coalitions in support of both a and b could be assembled as late as the deadline and so no strongest coalition would exist. Excluding such cases will be the role of the first condition used in our main result. And the second will require that the alternative whose support is strongest be in the core of the relevant coalitional game, so that any attempt to preempt its selection can be blocked.

It is instructive to note that throughout the above discussion we have had no need to specify the fine details of the bargaining procedure followed by the trade representatives. As already stated, our model will use a simultaneous voting protocol and our concern will be with dominance solvability. But these choices are driven mainly by expediency: This combination of elements happens to succeed in capturing the informal logic we have used to predict the outcome of the trade scenario. In this respect our theory has the flavor of an incentive compatibility analysis, and does not stand or fall on the "realism" of the extensive form employed.[5]

We proceed now by outlining our general bargaining model in Section 2, establishing the dominance solvability result in Section 3, considering two specialized environments (namely, binary choice and bilateral surplus division) in Section 4, and concluding with retrospective discussion of the theory in Section 5.

[5]Other models that are in some sense "procedure free" include those studied by Perry and Reny [19], Sakovics [21], Abreu and Gul [1], and Smith and Stacchetti [23].

2 Model

2.1 Extensive form

A group I of agents is jointly in charge of selecting exactly one alternative from a menu A. (Assume that I and A are both finite sets containing at least two elements.) The interval available for negotiation begins at date 0 and ends at a deadline normalized to date 1. If no alternative is selected at or before this date, then a situation of default arises indicated by the letter ω (where $\omega \notin A$). The full set of possible outcomes can then be defined as $X := A \times [0,1] \cup \{\langle \omega, 1\rangle\}$, with each combining a substantive result and a time value.

The agents are assumed to have preferences over the outcomes representable by utility functions $\langle u_i \rangle_{i\in I}$ with the following regularity properties.

Assumption 2.1. [**A**] For each $i \in I$, $a, b \in A$, and $t \in [0,1]$, we have $u_i(a,t) > u_i(b,t)$ if and only if $u_i(a,0) > u_i(b,0)$. [**B**] For each $i \in I$ and $a \in A$, the function $u_i(a,\cdot)$ on $[0,1]$ is strictly decreasing and continuous. [**C**] For each $i \in I$ we have $u_i(\omega,1) = 0$.

Hence attitudes towards the alternatives are stable (A), time is valuable and the costs associated with delay accrue in small increments (B), and default utilities are normalized to zero (C). Moreover, it follows that each mapping $u_i(a,\cdot)$ of $[0,1]$ onto $[u_i(a,1), u_i(a,0)]$ has a strictly decreasing inverse $u_i(a,\cdot)^{-1}$.

Although the bargainers could in principle reach an agreement at any instant $t \in [0,1]$, we shall restrict this possibility to a finite sequence $\langle k\Delta\rangle_{k\in K}$, where $K := \{0,1,\ldots,n\}$, of $n+1$ evenly-spaced *decision points.* (Thus $\Delta := 1/n$.) By increasing the discretization parameter $n \geq 1$, we can then approximate the underlying continuous time variable to any desired degree of precision.

Bargaining will take the form of repeated, simultaneous voting by the agents at the decision points. Each may vote for just one alternative at a time, and has also an option to abstain (or vote for "continuation") indicated by the letter β. We write $V := A \cup \{\beta\}$ for the action set and assume that at each decision point the actions taken at earlier decision points are observed by all agents.

The conclusion of an agreement is governed by an exogenous *decision rule* $\rho : \times_{i\in I} V \to V$. Writing v for the vector of votes by the agents at (some history at) decision point $k\Delta$, the rule ρ determines what happens after these votes are cast: Specifically, if $\rho(v) \in A$ then this alternative is agreed upon immediately and the game ends; if $\rho(v) = \beta$ and $k < n$ then voting commences again at decision point $[k+1]\Delta$; and if $\rho(v) = \beta$ and $k = n$ then the default outcome is realized. We impose two natural restrictions on the decision rule.

Assumption 2.2. [**A**] For each $i \in I$, $a \in A$, $v \in \times_{i\in I} V$, and $\overline{v}_i \in V$, if $\rho(v) = a$ and $v_i \neq a$ then $\rho(\overline{v}_i, v_{-i}) = a$. [**B**] For each $a \in A$ we have $\rho(a_I) = a$.[6]

Hence whether or not a particular alternative is implemented depends only on the agents voting for it (A, with $\overline{v}_i \neq a$), agreement is monotonic in the set of such agents (A, with

[6]Notation: Given $J \subset I$, we write $z_J = \langle z \rangle_{i\in J}$ and $z_{-J} = \langle z \rangle_{i\in I\setminus J}$.

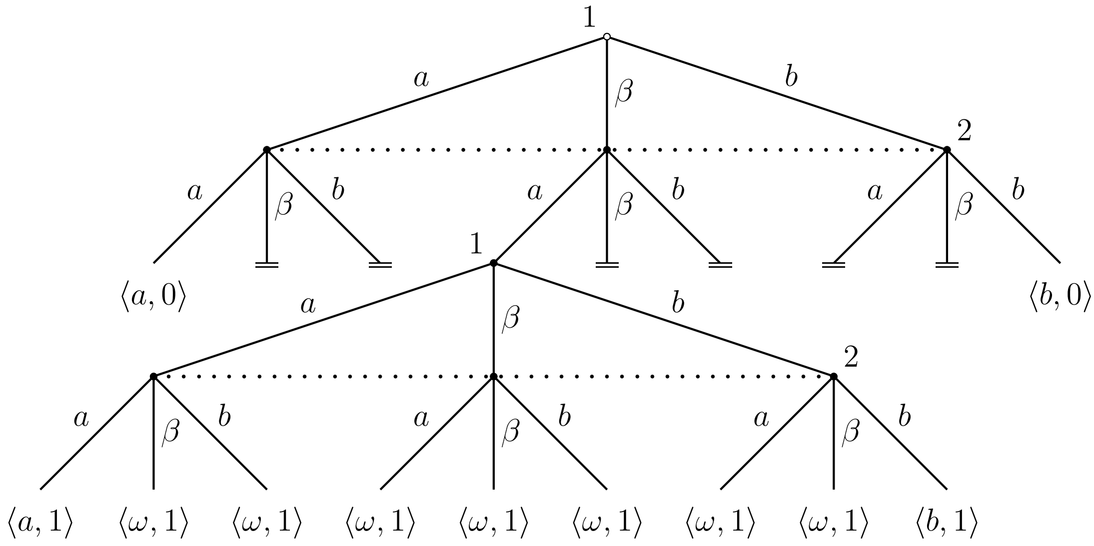

Figure 1: Part of an extensive form representation of our repeated, simultaneous voting game for the case of two agents (1 and 2), two alternatives (a and b), two decision points (0 and 1), and the unanimity decision rule. The abstention option is indicated by β and the default outcome by $\langle \omega, 1\rangle$. Note that outcomes (rather than the associated payoffs) are attached to the terminal histories and that at six non-terminal histories the tree has been truncated due to a shortage of space.

$\overline{v}_i = a$), and unanimous consent is sufficient to select any available option (B).[7]

For the case of $I = \{1, 2\}$, $A = \{a, b\}$, $n = 1$, and the unanimity rule, a portion of a game tree representing the procedure outlined above is illustrated in Figure 1.

For each $k \in K \setminus \{0\}$, let us write θ_k for the set of non-terminal histories at decision point $k\Delta$. Denoting by h^0 the null history and writing $\theta_0 := \{h^0\}$, we can proceed to define the full set $\Theta := \bigcup_{k \in K} \theta_k$ of non-terminal histories.[8] The subgame proceeding from an arbitrary $h \in \Theta$ will be indicated by Γ_h.

A strategy for player i is a function $s_i : \Theta \to V$, while a strategy profile is a vector $s = \langle s_i \rangle_{i \in I}$. Explicit construction of the mapping $s \mapsto \phi(s) \in X$ from strategy profiles to the resulting outcomes is routine but tedious, and is omitted here.

2.2 Solution concept

The idea that a weakly dominated strategy can be disregarded and effectively eliminated from a game has its roots in the normative analysis of statistical decision problems (see,

[7]Note that in addition to anonymous requirements for agreement ranging from a simple majority to unanimity, the decision rule can incorporate weighted or multiple majority requirements such as those used by the EU, as well as individual or joint vetoes such as those available in the UN Security Council.

[8]For example, in the case depicted in Figure 1 we have $\Theta = \{h^0, a\beta, ab, \beta a, \beta\beta, \beta b, ba, b\beta\}$.

e.g., Blackwell and Girshick [4]).[9] Repeated application of this idea has been described by Brandenburger and Keisler [5] as a "powerful" yet "conceptually puzzling" procedure, and as a result has generated both "widespread and fruitful applications" (Ewerhart [9]) and penetrating theoretical investigations of its epistemic basis.

Whatever its merits, iterative weak dominance does suffer from the practical drawback that the output of the procedure can depend upon the order in which strategies are deleted. Fortunately, a result due to Gretlein [12, p. 113] (see also Marx and Swinkels [15]) serves to mitigate this problem.

> [A]s long as players have strict preferences over the [outcomes] (of which there are a finite number), if they successively eliminate some subset of dominated strategies, ... then the set of outcomes not eliminated will be the same no matter ... which dominated strategies [are] eliminated at each stage.

Thus we can guarantee order-independence by prohibiting indifference on the part of the bargainers.

Assumption 2.3. For each $i \in I$, the restriction of u_i to $[A \times \{k\Delta\}_{k\in K}] \cup \{\langle\omega, 1\rangle\}$ (the finite domain of realizable outcomes) is one-to-one.[10]

Given $h \in \Theta$, let us write Ψ_h for the set of *outcomes* resulting from strategy profiles that survive iterative weak dominance in the subgame Γ_h proceeding from history h. For $k \in K$ we may also define $\Psi(k) := \cup_{h\in\theta_k}\Psi_h$, the outcomes that survive iterative weak dominance in at least one of the subgames beginning at decision point $k\Delta$. And the overall game will then be dominance solvable whenever $\exists\psi_0 \in X$ such that $\Psi(0) = \Psi_{h^0} = \{\psi_0\}$.

In view of Assumption 2.3 and Gretlein's theorem above, we can demonstrate that our game is dominance solvable by exhibiting any particular strategy elimination procedure that leaves a set of profiles all leading to the same outcome. The procedure that we shall employ can be understood intuitively as one of backward induction allowing for multiple rounds of deletion at each decision point. Starting from the final point, namely $n\Delta = 1$, we shall show first that (under specified conditions) $\exists\psi_n \in X$ such that $\Psi(n) = \{\psi_n\}$. We can then use ψ_n as the "continuation outcome" for each subgame proceeding from a history in θ_{n-1}, and can go on to show that (under the relevant conditions) $\exists\psi_{n-1} \in X$ such that $\Psi(n-1) = \{\psi_{n-1}\}$. This backward induction procedure will eventually bring us to the desired conclusion that $\exists\psi_0 \in X$ such that $\Psi(0) = \{\psi_0\}$, and we will then have shown that our game is dominance solvable with "solution" outcome ψ_0.[11]

[9] A strategy s_i weakly dominates another strategy $\hat{s}_i$ for agent i if for each profile $\overline{s}_{-i}$ of strategies for $I\backslash\{i\}$ we have $u_i(\phi(s_i, \overline{s}_{-i})) \geq u_i(\phi(\hat{s}_i, \overline{s}_{-i}))$ and for some such $\overline{s}_{-i}$ we have $u_i(\phi(s_i, \overline{s}_{-i})) > u_i(\phi(\hat{s}_i, \overline{s}_{-i}))$.

[10] Observe that this rules out both "genuine" indifference (of the form $u_i(a, 0) = u_i(b, 0)$ for $a \neq b$) and "coincidental" indifference (of either the form $u_i(a, k\Delta) = 0$ or the form $u_i(a, k\Delta) = u_i(b, m\Delta)$ for $a \neq b$ and $k < m$). While the second prohibition is not a major restriction in view of our discretization of time, the first has significant economic content.

[11] Our use of the term "backward induction" here and throughout the paper is meant to be suggestive rather than literal: Since our game has simultaneous moves and hence imperfect information, ordinary backward induction is not well-defined. Note also that our procedure for eliminating strategies resembles the application of iterated conditional dominance (see, e.g., Fudenberg and Tirole [10, pp. 128–129]), the main difference being that we inspect for conditional *weak* dominance at each iteration and rely on Gretlein's theorem for order-independence of the resulting outcome set.

In accordance with our interpretation of ψ_{k+1} as the continuation outcome for the subgames proceeding from histories in θ_k, let us write $\Psi(n+1) = \{\psi_{n+1}\} := \{\langle\omega, 1\rangle\}$ to reflect that the last decision point is the deadline for agreement.

The notion of dominance solvability outlined above resembles that used by Moulin [16] in his well-known analysis of voting games. (See also Gretlein [11].) As defined the two concepts are distinct: While Moulin asks only that the strategies remaining for each player after iterative deletion be equivalent, we make the stronger demand that the remaining strategy profiles all lead to the same outcome. In the presence of Assumption 2.3, however, these requirements coincide and our game will (under specified conditions) be "d-solvable" [16, p. 1339].

3 Analysis

3.1 Backward induction lemma

In this section we define a notion of collective acceptability of an alternative at a decision point with respect to a continuation outcome, and then use this definition to state and prove the backward induction lemma upon which our analysis is based.

The following will be the relevant notion of acceptability.

Definition 3.1. Given $k \in K$ and $x \in X$, an alternative a is said to be *viable at $k\Delta$ with respect to x* if $\exists J \subset I$ such that $\rho(a_J, \beta_{-J}) = a$ and $\forall i \in J$ we have $u_i(a, k\Delta) \geq u_i(x)$.

In other words, an alternative is viable at a decision point with respect to an outcome — to be thought of as the consequences of continuation — if the members of some coalition with the ability to implement this alternative would be willing to do so were they each to conclude that no other agreement could at present be reached. Note that it is possible for zero, one, or more than one alternative to have this property at the same time and with respect to the same outcome.

When we put the above definition to use at a given decision point $k\Delta$, the role of the continuation outcome will be filled by the (common) iterative dominance solution ψ_{k+1} of the subgames proceeding from histories in θ_{k+1}. To apply the concept of viability in this fashion we shall of course first have to confirm that such a solution exists; i.e., that $\exists\psi_{k+1} \in X$ such that $\Psi(k+1) = \{\psi_{k+1}\}$. But since our method will be one of backward induction, this will already have been shown by the time we come to consider decision point $k\Delta$.

Our lemma states that if $\Psi(k+1)$ contains a single outcome, then $\Psi(k)$ also contains a single outcome provided that no more than one alternative is viable at $k\Delta$ with respect to ψ_{k+1}. If no alternative is viable then iterative weak dominance selects the continuation outcome, whereas if exactly one alternative is viable then this alternative is agreed upon at once.

Lemma 3.2. *Given $k \in K$, suppose that $\exists\psi_{k+1} \in X$ such that $\Psi(k+1) = \{\psi_{k+1}\}$.* **[A]** *If no alternative is viable at $k\Delta$ with respect to ψ_{k+1}, then $\Psi(k) = \Psi(k+1) = \{\psi_{k+1}\}$.* **[B]** *If $a^\circ \in A$ is uniquely viable at $k\Delta$ with respect to ψ_{k+1}, then $\Psi(k) = \{\langle a^\circ, k\Delta\rangle\}$.*

Proof. Given $i \in I$, $a \in A$, $h \in \theta_k$, and a profile $\overline{s}_{-i}$ of strategies for $I \setminus \{i\}$, call player i "pivotal" for a at h against $\overline{s}_{-i}$ whenever both $\rho(a, \overline{s}_{-i}(h)) = a$ and $\rho(\beta, \overline{s}_{-i}(h)) \neq a$, in which case Assumption 2.2A implies that $\rho(\beta, \overline{s}_{-i}(h)) = \beta$.

Given $i \in I$, $a \in A$, and $h \in \theta_k$, consider any strategy s_i for i satisfying $s_i(h) = a$. Define a new strategy $\hat{s}_i$ by $\hat{s}_i(h) = \beta$ and $\hat{s}_i(h') = s_i(h')$ for each $h' \in \Theta \setminus \{h\}$. Restricting these two strategies to the subgame Γ_h, we find that they yield the same outcome $\forall \overline{s}_{-i}$ against which i is not pivotal for a at h. Moreover, $\forall \overline{s}_{-i}$ against which i *is* pivotal, s_i yields $\langle a, k\Delta \rangle$ while $\hat{s}_i$ yields ψ_{k+1} once all strategies that are iteratively dominated in proper subgames of Γ_h have been deleted. And it follows that s_i is iteratively dominated in Γ_h whenever $u_i(\psi_{k+1}) > u_i(a, k\Delta)$.

[A] Suppose that a given $a \in A$ is not viable at $k\Delta$ with respect to ψ_{k+1}, and consider any $J \subset I$ such that $\rho(a_J, \beta_{-J}) = a$. Then $\exists i \in J$ such that $u_i(\psi_{k+1}) > u_i(a, k\Delta)$, and so $\forall h \in \theta_k$ the restriction to Γ_h of any strategy s_i for i satisfying $s_i(h) = a$ is iteratively dominated. But this implies that $\langle a, k\Delta \rangle$ cannot survive iterative dominance in Γ_h, and therefore if no alternative is viable at $k\Delta$ with respect to ψ_{k+1} we have that $\Psi_h = \{\psi_{k+1}\}$. Finally, since $h \in \theta_k$ is arbitrary, we can conclude that $\Psi(k) = \{\psi_{k+1}\}$ as desired.

[B] Since a° is uniquely viable at $k\Delta$ with respect to ψ_{k+1}, no $b \in A \setminus \{a^\circ\}$ can be viable and so $\forall h \in \theta_k$ the corresponding outcomes $\langle b, k\Delta \rangle$ cannot survive iterative dominance in Γ_h.

Given $i \in I$, $v_i \in V \setminus \{a^\circ\}$, and $h \in \theta_k$, consider any strategy s_i for i with $s_i(h) = v_i$. Define a new strategy $\hat{s}_i$ by $\hat{s}_i(h) = a^\circ$ and $\hat{s}_i(h') = s_i(h')$ for each $h' \in \Theta \setminus \{h\}$. Restricting these two strategies to the subgame Γ_h, we find that (after the outcomes $\langle b, k\Delta \rangle$ for $b \neq a^\circ$ have been ruled out) they yield the same result $\forall \overline{s}_{-i}$ against which i is not pivotal for a° at h. Moreover, $\forall \overline{s}_{-i}$ against which i *is* pivotal, s_i yields ψ_{k+1} while $\hat{s}_i$ yields $\langle a^\circ, k\Delta \rangle$ once all strategies that are iteratively dominated in proper subgames of Γ_h have been deleted. And it follows that s_i is iteratively dominated in Γ_h whenever $u_i(a^\circ, k\Delta) > u_i(\psi_{k+1})$.

Since a° is viable at $k\Delta$ with respect to ψ_{k+1}, $\exists J \subset I$ such that $\rho(a^\circ_J, \beta_{-J}) = a^\circ$ and $\forall i \in J$ we have $u_i(a^\circ, k\Delta) \geq u_i(\psi_{k+1})$. For each such i it follows that $u_i(a^\circ, k\Delta) > u_i(\psi_{k+1})$ by Assumption 2.3. From this we can conclude that $\forall h \in \theta_k$ the restriction to Γ_h of any strategy s_i for $i \in J$ satisfying $s_i(h) \neq a^\circ$ is iteratively dominated, leaving for such i only strategies that specify a° at h. Regardless of the profile $\overline{s}_{-J}$ of strategies for $I \setminus J$, we have that $\rho(a^\circ_J, \overline{s}_{-J}(h)) = a^\circ$ by Assumption 2.2A, and therefore $\Psi_h = \{\langle a^\circ, k\Delta \rangle\}$. Finally, since $h \in \theta_k$ is arbitrary, we can conclude that $\Psi(k) = \{\langle a^\circ, k\Delta \rangle\}$ as desired. □

3.2 The consensus point

We now turn our attention to a concept at the heart of the present theory: that of the "consensus point" associated with the bargaining problem under consideration. What we aim to show is that this point is the de facto deadline for an agreement, even though it may precede the official deadline (which we have normalized to date 1).

Our definition of the consensus point is by way of two useful prior concepts.

Definition 3.3. **[A]** The *latest acceptance point* of agent i for alternative a is defined by $\xi_i(a) := \sup\{t \in [0,1] : u_i(a,t) \geq 0\}$. **[B]** The *latest feasible point* of alternative a

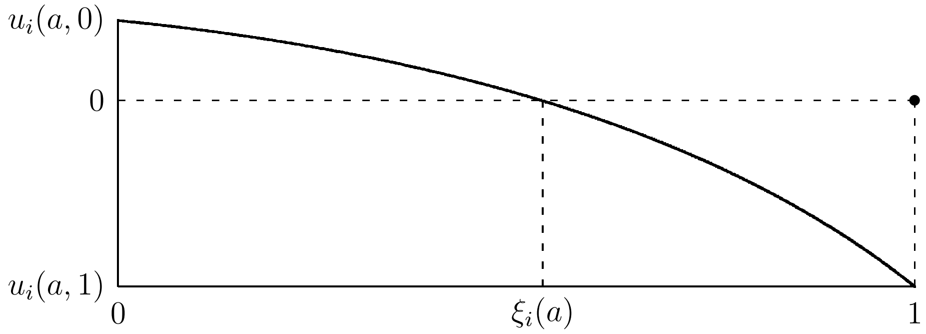

Figure 2: The "latest acceptance point" $\xi_i(a)$. When $u_i(a,0) \geq 0$, this is the last moment at which agent i weakly prefers agreement on alternative a to eventual default. When $u_i(a,0) < 0$ (in which case no such last moment exists), we have $\xi_i(a) = -\infty$.

is defined by $\Xi(a) := \max\{\min_{i\in J} \xi_i(a) : J \subset I \text{ and } \rho(a_J, \beta_{-J}) = a\}$. **[C]** The *consensus point* is defined by $\Xi^* := \max_{a\in A} \Xi(a)$.

The latest acceptance point $\xi_i(a)$ measures the appeal of alternative a to agent i on the time dimension of the negotiation. When $u_i(a,1) \geq 0$ this point takes on the value 1; when $u_i(a,0) \geq 0 > u_i(a,1)$ it takes on the value $u_i(a,\cdot)^{-1}(0)$; and when $u_i(a,0) < 0$ it takes on the value $-\infty$. (See Figure 2.) Thus the agent can do no more to demonstrate his delight at the prospect of alternative a than to exhibit $\xi_i(a) = 1$, and no more to demonstrate his displeasure than to exhibit $\xi_i(a) = -\infty$.

Similarly, the latest feasible point $\Xi(a)$ measures the appeal of alternative a to the collective, equalling the acceptance point of the *most skeptical* member of that coalition with the ability to implement a that is *easiest* to assemble. The consensus point is then simply the last of the alternatives' latest feasible points, with $\Xi^* = -\infty$ indicating that $\forall a \in A$ we have $\Xi(a) = -\infty$ and $\Xi^* = 1$ indicating that $\exists a \in A$ such that $\Xi(a) = 1$.

Example 3.4. Let $I = \{1,2,3\}$, $A = \{a_1, a_2, a_3\}$, and

$$\begin{bmatrix} u_1(a_1,t) & u_1(a_2,t) & u_1(a_3,t) \\ u_2(a_1,t) & u_2(a_2,t) & u_2(a_3,t) \\ u_3(a_1,t) & u_3(a_2,t) & u_3(a_3,t) \end{bmatrix} = \begin{bmatrix} 6/5-t & -1/5-t & -2/5-t \\ 3/5-t & 4/5-t & -3/5-t \\ 1/5-t & 2/5-t & 7/5-t \end{bmatrix}.$$

For each $a \in A$ and $v \in \times_{i\in I} V$, let $\rho(v) = a$ if and only if $|\{i \in I : v_i = a\}| \geq 2$. We then have the latest acceptance points

$$\begin{bmatrix} \xi_1(a_1) & \xi_1(a_2) & \xi_1(a_3) \\ \xi_2(a_1) & \xi_2(a_2) & \xi_2(a_3) \\ \xi_3(a_1) & \xi_3(a_2) & \xi_3(a_3) \end{bmatrix} = \begin{bmatrix} 1 & -\infty & -\infty \\ 3/5 & 4/5 & -\infty \\ 1/5 & 2/5 & 1 \end{bmatrix},$$

the latest feasible points $\langle \Xi(a_1), \Xi(a_2), \Xi(a_3)\rangle = \langle 3/5, 2/5, -\infty\rangle$, and the consensus point $\Xi^* = 3/5$.

Example 3.5. Let the agents, alternatives, and preferences be as in Example 3.4. For each $a \in A$ and $v \in \times_{i \in I} V$, let $\rho(v) = a$ if and only if both $|\{i \in I : v_i = a\}| \geq 2$ and $v_3 = a$. We then have the latest feasible points $\langle \Xi(a_1), \Xi(a_2), \Xi(a_3) \rangle = \langle 1/5, 2/5, -\infty \rangle$ and the consensus point $\Xi^* = 2/5$.

Our interest in the consensus point is due to its status as the earliest date after which we can guarantee that there will be no agreement. When $\Xi^* < 0$ (i.e., when $\Xi^* = -\infty$) this implies that there can never be agreement at any decision point; when $\Xi^* = 1$ the conclusion is true but vacuous; and when $0 \leq \Xi^* < 1$ it tells us that while agreement may be possible at the outset, any such possibility will vanish if the game lasts long enough.

The necessity of disagreement after the consensus point can be shown inductively. By construction, no alternative can be viable with respect to the default outcome $\langle \omega, 1 \rangle$ at any decision point $k\Delta \in (\Xi^*, 1]$. But the default outcome is precisely what will occur if the deadline is reached without an agreement, and the result then follows easily from Lemma 3.2A.

Proposition 3.6. *For each $k \in K$, if $k\Delta > \Xi^*$ then $\Psi(k) = \{\langle \omega, 1 \rangle\}$.*

Proof. The proof is by induction. Firstly, recall that $\Psi(n+1) = \{\langle \omega, 1 \rangle\}$. Now, given $k \in K$ such that $k\Delta > \Xi^*$, suppose that $\Psi(k+1) = \{\langle \omega, 1 \rangle\}$. If some $a \in A$ were viable at $k\Delta$ with respect to $\langle \omega, 1 \rangle$, then by definition there would exist a $J \subset I$ such that $\rho(a_J, \beta_{-J}) = a$ and $\forall i \in J$ we have $u_i(a, k\Delta) \geq 0$. For each such i we would then have $\xi_i(a) \geq k\Delta$, which would imply that $\Xi^* \geq \Xi(a) \geq \min_{i \in J} \xi_i(a) \geq k\Delta$, a contradiction. Thus no alternative is viable at $k\Delta$ with respect to $\langle \omega, 1 \rangle$, so by Lemma 3.2A we have that $\Psi(k) = \Psi(k+1) = \{\langle \omega, 1 \rangle\}$ as desired. $\square$

3.3 Agreement at the consensus point

Having shown (in Proposition 3.6) that no agreement will be reached after the consensus point, we now give a condition under which an agreement *will* be reached *at* (or rather, because of our discretization of time, *near*) this date.

The condition to be used requires simply that no more than one alternative be viable at the deadline with respect to the default outcome.

Condition 3.7 (Terminal Solvability)**.** At most one alternative is viable at $n\Delta = 1$ with respect to $\psi_{n+1} = \langle \omega, 1 \rangle$.

This requirement is severe: Many situations of interest (e.g., the classic surplus-division problem with zero payoff from disagreement) will admit coalitions that could form in support of two or more distinct alternatives at the deadline, thus violating the condition. But the failure of our theory to make predictions about such situations has nothing to do with their *dynamic* structure, since at the deadline any bargaining problem is by definition "static" — having no remaining intertemporal aspect. We deliberately refrain here from taking a position as to the resolution of problems of this sort, and so it is only right that situations which call for a theory of static bargaining should remain beyond the scope of the present analysis. (See Section 5.2 for further discussion of this point.)

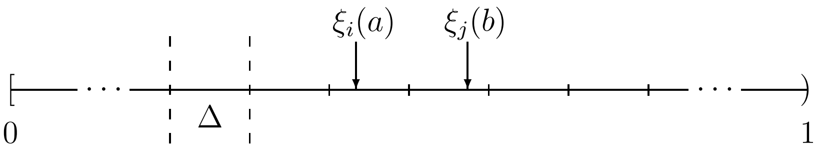

Figure 3: Assumptions 3.8 and 3.10. Any two distinct latest acceptance points $\xi_i(a)$ and $\xi_j(b)$ that fall in the interval $[0, 1)$ must take on different values (a genericity requirement) and must be separated by a decision point (a "fineness" requirement).

When Terminal Solvability holds, dominance solvability of the subgames proceeding from histories in θ_n is guaranteed by Lemma 3.2. If an alternative is viable at $n\Delta$ with respect to ψ_{n+1}, then the consensus point is the deadline and we have that the alternative in question will be agreed upon at this date. If, on the other hand, no alternative is viable, then what we must show is that (when $\Xi^* \geq 0$) an agreement will be reached at the latest decision point $k^*\Delta := [\sup\{k \in K : k\Delta \leq \Xi^*\}]\Delta$ no later than the consensus point. To establish this we will need a genericity assumption ensuring that as we proceed backwards along the interval $[0, 1)$, two alternatives do not become viable with respect to the default outcome at *precisely the same instant*. (See Figure 3.)

Assumption 3.8. For each $i, j \in I$ and $a, b \in A$ such that both $u_i(a, 0) \geq 0 > u_i(a, 1)$ and $u_j(b, 0) \geq 0 > u_j(b, 1)$, we have $u_i(a, \cdot)^{-1}(0) = u_j(b, \cdot)^{-1}(0)$ only if both $i = j$ and $a = b$.[12]

Together with the above condition, this assumption allows us (when $\Xi^* \geq 0$) to identify a unique alternative whose latest feasible point is the consensus point.

Lemma 3.9. *Let Terminal Solvability hold. If* $\Xi^* \geq 0$ *then there exists a unique* $a^* \in A$ *such that* $\Xi(a^*) = \Xi^*$.

Proof. Suppose that there exist distinct $a, b \in A$ such that $\Xi(a) = \Xi(b) = \Xi^*$. In the event that $\Xi^* = 1$ both a and b would be viable at $n\Delta = 1$ with respect to $\psi_{n+1} = \langle \omega, 1 \rangle$, which is ruled out by Terminal Solvability. In the event that $0 \leq \Xi^* < 1$ there would exist $i, j \in I$ such that $u_i(a, \cdot)^{-1}(0) = \xi_i(a) = \Xi(a) = \Xi(b) = \xi_j(b) = u_j(b, \cdot)^{-1}(0)$, which is forbidden by Assumption 3.8. Hence $\Xi^* < 0$, and the result follows by contraposition. $\square$

When applicable, this result will serve as the definition of a^*, the first alternative to become viable with respect to the default outcome as we proceed backwards from the deadline towards the beginning of the negotiation.

One further technical assumption is needed to show that a^* will be agreed upon at $k^*\Delta$ whenever $\Xi^* \geq 0$. Not only must there be a unique alternative whose latest feasible point is the consensus point; the discretization of time must be fine enough to separate this point from the latest feasible points of the other alternatives. (Again see Figure 3.)

[12]This assumption is somewhat stronger than necessary: What is actually required is that among the alternatives with latest feasible points in $[0, 1)$, that with the last such point is unique. Note also that we permit the latest acceptance points of the agents for the alternatives to coincide at $-\infty$ and 1, since these are the extrema of the set of values that any particular $\xi_i(a)$ can take on.

Assumption 3.10. For each $i, j \in I$ and $a, b \in A$ such that both $u_i(a, 0) \geq 0 > u_i(a, 1)$ and $u_j(b, 0) \geq 0 > u_j(b, 1)$, if either $i \neq j$ or $a \neq b$ then we have that $\Delta < |u_i(a, \cdot)^{-1}(0) - u_j(b, \cdot)^{-1}(0)|$.[13]

And with this assumption in place, we obtain the desired result.

Proposition 3.11. *Let Terminal Solvability hold. If* $\Xi^* \geq 0$ *then* $\Psi(k^*) = \{\langle a^*, k^*\Delta\rangle\}$.

Proof. If $\Xi^* \geq 0$ then $\{k \in K : k\Delta \leq \Xi^*\} \neq \emptyset$ and so $k^* \in K$. By the definition of k^* we then have $[k^*+1]\Delta > \Xi^*$, so $\Psi(k^*+1) = \{\langle \omega, 1\rangle\}$ by Proposition 3.6. Since $\Xi(a^*) = \Xi^* \geq k^*\Delta$, alternative a^* is viable at $k^*\Delta$ with respect to $\langle \omega, 1\rangle$. Moreover, if some $a \neq a^*$ were also viable then by Lemma 3.9 we would have $k^*\Delta \leq \Xi(a) < \Xi(a^*) = \Xi^* < [k^* + 1]\Delta$, contradicting Assumption 3.10. Hence a^* is uniquely viable at $k^*\Delta$ with respect to $\langle \omega, 1\rangle$, and by Lemma 3.2B it follows that $\Psi(k^*) = \{\langle a^*, k^*\Delta\rangle\}$ as desired. □

3.4 Agreement at the outset

In this section we conclude the analysis of our bargaining game by providing a condition under which any agreement will be immediate.

When $\Xi^* \geq 0$, Proposition 3.11 ensures that under Terminal Solvability alternative a^* (defined by Lemma 3.9) will be agreed upon at decision point $k^*\Delta$. If, in addition, $k^* > 0$, then since time is valuable we have that a^* is viable at the previous decision point $[k^* - 1]\Delta$ with respect to the continuation outcome $\psi_{k^*} = \langle a^*, k^*\Delta\rangle$. If no *other* alternative were viable, Lemma 3.2B would secure agreement at the earlier point and hence, by induction, at the outset. This can be achieved by requiring two things: Firstly, we ask that a^* be in the core of the coalitional game associated with the bargaining problem under consideration.

Condition 3.12 (Core Membership)**.** If $\Xi^* \geq 0$ then for each $a \in A$ such that $\Xi(a) = \Xi^*$ and each $b \in A \setminus \{a\}$ and $J \subset I$ such that $\rho(b_J, \beta_{-J}) = b$, there exists an $i \in J$ such that $u_i(a, 0) > u_i(b, 0)$.

Example 3.13. In Example 3.4 we have $a^* = a_1$; $\rho(\beta, a_2, a_2) = a_2$; $u_2(a_1, 0) < u_2(a_2, 0)$; and $u_3(a_1, 0) < u_3(a_2, 0)$; and hence Core Membership fails.

Example 3.14. In Example 3.5 we have $a^* = a_2$; $\rho(v) = a_1$ only if $v_3 = a_1$ and, moreover, $u_3(a_2, 0) > u_3(a_1, 0)$; and $\rho(v) = a_3$ only if $|\{i \in I : v_i = a_3\}| \geq 2$, $u_1(a_2, 0) > u_1(a_3, 0)$, and $u_2(a_2, 0) > u_2(a_3, 0)$; and hence Core Membership holds.

And secondly, we ask that the time interval be sufficiently small relative to the variability of the agents' utility functions, which is to say that n is sufficiently large. (See Figure 4.)

Assumption 3.15. For each $i \in I$, each $a, b \in A$ such that $u_i(a, 0) > u_i(b, 0)$, and each $t \in [0, 1 - \Delta]$, we have $u_i(a, t + \Delta) > u_i(b, t)$.

[13]It is important to realize that both Assumption 3.10 and Assumption 3.15 below require only that the parameter n be larger than some given finite value. These assumptions do not relate to any sort of limit game in which the discretization of time is somehow infinitely fine. Indeed, nowhere in this paper are there limiting arguments or results of any kind.

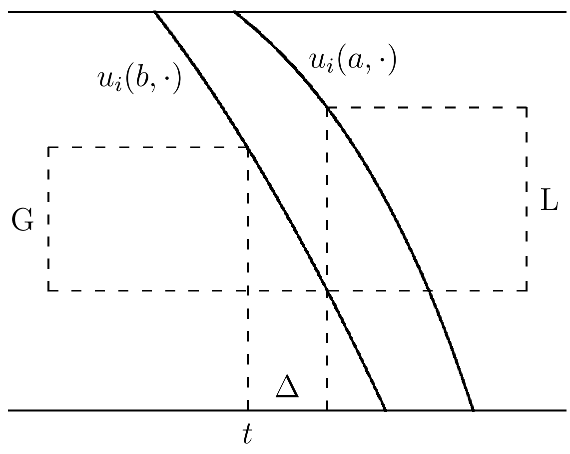

Figure 4: Assumption 3.15. The discretization parameter n must be large enough that any agent i's loss (L) from replacing a more attractive alternative a with a less attractive alternative b is never outweighed by his gain (G) from concluding an agreement $\Delta = 1/n$ units of time earlier.

We are now in a position to state and prove our main result.

Theorem 3.16. *Let both Terminal Solvability and Core Membership hold. Then* $\Xi^* \geq 0$ *only if* $\Psi(0) = \{\langle a^*, 0\rangle\}$*, whereas* $\Xi^* < 0$ *only if* $\Psi(0) = \{\langle \omega, 1\rangle\}$.

Proof. For the case of $\Xi^* \geq 0$, the proof is by induction. Firstly, by Proposition 3.11 we have that $\Psi(k^*) = \{\langle a^*, k^*\Delta\rangle\}$. Now, given $k \in K$ such that $k < k^*$, suppose that $\Psi(k+1) = \{\langle a^*, [k+1]\Delta\rangle\}$. Since $\rho(a_I^*) = a^*$ by Assumption 2.2B and $\forall i \in I$ we have $u_i(a^*, k\Delta) > u_i(a^*, [k+1]\Delta)$ by Assumption 2.1B, alternative a^* is viable at $k\Delta$ with respect to $\langle a^*, [k+1]\Delta\rangle$. Given $a \in A \setminus \{a^*\}$ and $J \subset I$ such that $\rho(a_J, \beta_{-J}) = a$, Core Membership guarantees the existence of an $i \in J$ such that $u_i(a^*, 0) > u_i(a, 0)$, and hence $u_i(a^*, [k+1]\Delta) > u_i(a, k\Delta)$ by Assumption 3.15. It follows that a is not viable and hence a^* is uniquely viable at $k\Delta$ with respect to $\langle a^*, [k+1]\Delta\rangle$, so that $\Psi(k) = \{\langle a^*, k\Delta\rangle\}$ by Lemma 3.2B.

For the case of $\Xi^* < 0$, the conclusion follows from Proposition 3.6. □

In other words: Terminal Solvability and Core Membership together guarantee dominance solvability. Nonnegativity of the consensus point is the criterion for agreement. When agreement occurs it is immediate, and the alternative agreed upon is that whose latest acceptance point equals the consensus point.[14]

[14] To emphasize what is no doubt obvious by this stage, neither Theorem 3.16 nor any other result in the present paper has anything to say about the *equilibrium* outcomes of our bargaining game (at least not directly). These will typically be very numerous: For example, if no agent can unilaterally either force or prevent any alternative, then every feasible outcome will be supported by an equilibrium. Refinements such as subgame and trembling-hand perfection generally do little to reduce this multiplicity. But our theorem shows that in certain circumstances iterative weak dominance *can* make sharper predictions — and of course does so without any need for an assumption of equilibrium behavior in the first place.

4 Specialized environments

4.1 Binary choice

One class of bargaining environments to which Theorem 3.16 can be applied contains binary (two-alternative) choice problems in which, at the deadline, each agent ranks the default outcome strictly between the two available agreement outcomes. Here no compromise is possible between the two options, and at date 1 each bargainer would rather fail to reach an agreement than accept his disfavored alternative — though he may accept it earlier in order to avoid delay.[15]

Proposition 4.1. *Let $A = \{0, 1\}$. If for each $i \in I$ we have either $u_i(0, 1) > 0 > u_i(1, 1)$ or $u_i(1, 1) > 0 > u_i(0, 1)$, then both Terminal Solvability and Core Membership hold.*

Proof. Given $a \in A$, let $F(a) := \{i \in I : u_i(a, 0) > u_i(1 - a, 0)\}$ and observe that then a is viable at $n\Delta = 1$ with respect to $\psi_{n+1} = \langle \omega, 1\rangle$ if and only if $\rho(a_{F(a)}, \beta_{F(1-a)}) = a$. It follows that a is viable only if $\rho(a_{F(a)}, [1 - a]_{F(1-a)}) = a$ (using Assumption 2.2A), which implies that $\rho(\beta_{F(a)}, [1 - a]_{F(1-a)}) \neq 1 - a$ and so $1 - a$ is *not* viable. Therefore Terminal Solvability holds.

If Core Membership fails, then $\exists a \in A$ such that $\Xi(a) = \Xi^* \geq \Xi(1 - a)$ and a $J \subset I$ such that both $\rho([1-a]_J, \beta_{-J}) = 1-a$ and $\forall i \in J$ we have $u_i(a, 0) \leq u_i(1-a, 0)$. It follows that $J \subset F(1 - a)$ and $\rho([1 - a]_{F(1-a)}, \beta_{F(a)}) = 1 - a$ by Assumptions 2.1A and 2.2A, and hence $1 - a$ is viable at $n\Delta = 1$ with respect to $\psi_{n+1} = \langle \omega, 1\rangle$. This, however, implies that $\Xi(a) \geq \Xi(1 - a) = 1$ and so a is also viable, contradicting Terminal Solvability. By contraposition, Core Membership must hold. $\square$

Example 4.2. Let $I = \{1, 2, 3, 4, 5\}$, $A = \{0, 1\}$, and

$$\begin{bmatrix} u_1(0,t) & u_1(1,t) \\ u_2(0,t) & u_2(1,t) \\ u_3(0,t) & u_3(1,t) \\ u_4(0,t) & u_4(1,t) \\ u_5(0,t) & u_5(1,t) \end{bmatrix} = \begin{bmatrix} 8/5 - t & -1/5 - t \\ 6/5 - t & 1/5 - t \\ 4/5 - t & 6/5 - t \\ 2/5 - t & 7/5 - t \\ -2/5 - t & 8/5 - t \end{bmatrix}.$$

For each $a \in A$ and $v \in \times_{i \in I} V$, let $\rho(v) = a$ if and only if $|\{i \in I : v_i = a\}| \geq 4$. We then have the latest acceptance points

$$\begin{bmatrix} \xi_1(0) & \xi_1(1) \\ \xi_2(0) & \xi_2(1) \\ \xi_3(0) & \xi_3(1) \\ \xi_4(0) & \xi_4(1) \\ \xi_5(0) & \xi_5(1) \end{bmatrix} = \begin{bmatrix} 1 & -\infty \\ 1 & 1/5 \\ 4/5 & 1 \\ 2/5 & 1 \\ -\infty & 1 \end{bmatrix},$$

the latest feasible points $\langle \Xi(0), \Xi(1)\rangle = \langle 2/5, 1/5\rangle$, and the consensus point $\Xi^* = 2/5$. Note that here the chosen alternative $a^* = 0$ is favored by a minority of the agents.

[15] One example of such a problem might be that faced by a jury charged with deciding whether to acquit or convict, and given both rules for reaching a verdict (e.g., majority or unanimity) and an explicit or implicit deadline at which point it will be declared to be deadlocked (the default outcome).

Example 4.3. Let the agents, alternatives, and preferences be as in Example 4.2. For each $a \in A$ and $v \in \times_{i \in I} V$, let $\rho(v) = a$ if and only if $|\{i \in I : v_i = a\}| \geq 3$. We then have the latest feasible points $\langle \Xi(0), \Xi(1) \rangle = \langle 4/5, 1 \rangle$ and the consensus point $\Xi^* = 1$. Note that here the chosen alternative $a^* = 1$ is the one favored by agent 3, who functions as a sort of median voter.

4.2 Bilateral surplus division

As a second application, we now specialize our theory to the bilateral surplus division problem studied by Rubinstein [20], among others. In this setting we have two agents, a finite number of alternatives drawn from $A^\circ := \{\langle a_1, a_2 \rangle \in \Re^2_+ : a_1 + a_2 = 1\}$, and the unanimity rule.

Proposition 4.4. *Let $I = \{1, 2\}$ and $A \subset A^\circ$, and $\forall a \in A$ and $v \in \times_{i \in I} V$ let $\rho(v) = a$ if and only if $v = a_I$. Moreover, suppose that for each $i \in I$ there exists a constant γ_i such that $\forall a \in A$ we have $u_i(a, 1) = a_i - \gamma_i$. If $\gamma_1 + \gamma_2 \geq 1$, then both Terminal Solvability and Core Membership hold.*

Proof. If a given $a \in A$ is viable at $n\Delta = 1$ with respect to $\psi_{n+1} = \langle \omega, 1 \rangle$ then both $\gamma_1 \leq a_1$ and $\gamma_2 \leq a_2$ and hence $\gamma_1 + \gamma_2 = 1$ (since $a \in A^\circ$), which can be true only if $\gamma_1 = a_1$ and $\gamma_2 = a_2$. If some $b \in A$ is also viable, then similarly $\gamma_1 = b_1$ and $\gamma_2 = b_2$, and it follows that $b = a$. Therefore Terminal Solvability holds.

If Core Membership fails, then there exist distinct $a, b \in A$ such that $\Xi(a) = \Xi^* \geq 0$ and (in view of the unanimity rule) $\forall i \in I$ we have $u_i(a, 0) \leq u_i(b, 0)$. It follows that $\forall i \in I$ we have $\xi_i(b) \geq \xi_i(a)$, that $\Xi(b) = \min_{i \in I} \xi_i(b) \geq \min_{i \in I} \xi_i(a) = \Xi(a) = \Xi^*$, and hence that $\Xi(b) = \Xi^*$. But since $\Xi^* \geq 0$ and the alternatives a and b are distinct, Lemma 3.9 implies that this contradicts Terminal Solvability. By contraposition, Core Membership must hold. $\square$

Example 4.5. For each $a \in A$ let $u_1(a, t) = a_1 - t/5 - 2/5$ and $u_2(a, t) = a_2 - 2t/5 - 1/5$. We then have

$$\Xi(a) = \min\{\xi_1(a), \xi_2(a)\} = \begin{cases} \xi_1(a) = -\infty & \text{if } 0 \leq a_1 < 6/15, \\ \xi_1(a) = 5a_1 - 2 & \text{if } 6/15 \leq a_1 \leq 8/15, \\ \xi_2(a) = 2 - 5a_1/2 & \text{if } 8/15 < a_1 \leq 12/15, \\ \xi_2(a) = -\infty & \text{if } 12/15 < a_1 \leq 1, \end{cases}$$

(see Figure 5). And when $\forall a^\circ \in A^\circ$ there exists a nearby $a \in A$, we have also $\Xi^* \approx 2/3$ and immediate agreement on $a^* \approx \langle 8/15, 7/15 \rangle$.

Example 4.6. For each $a \in A$ let $u_1(a, t) = a_1 - t/6 - 1/2$ and $u_2(a, t) = a_2 - t/6 - 19/30$. We then have $\Xi(a) = \min\{\xi_1(a), \xi_2(a)\} = -\infty$ everywhere and eventual disagreement.

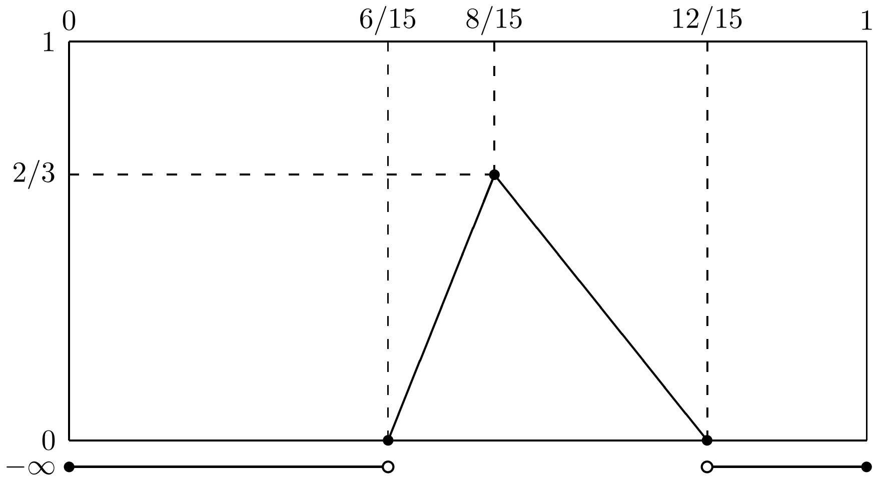

Figure 5: Example 4.5 — bilateral surplus division with constant delay costs. The allocation a_1 to agent 1 is measured on the horizontal axis and time on the vertical axis. Plotted is the latest feasible point function $\Xi(a)$ (over the domain A°), which reaches its maximum value of 2/3 at the allocation $\langle a_1, a_2 \rangle = \langle 8/15, 7/15 \rangle$.

5 Discussion

5.1 The horizon and the default outcome

Apart from the requirement (implicit in our solution concept) of complete information, the most conspicuous structural assumption of our model is that the bargaining will end at a fixed, known deadline. It would appear to be possible to relax this assumption and allow the negotiations to continue indefinitely, as in the Rubinstein model and its many variants. And in fact if we were to do so — replacing default with the outcome of endless deliberations — then the machinery of our analysis would, mutatis mutandis, continue to be perfectly well-defined. In particular, we could compute the consensus point Ξ^* just as before and could formulate appropriate versions of the conditions used in Theorem 3.16.

It is not true, however, that our results would continue to hold in this infinite-horizon setting. While Lemma 3.2 would remain unaffected, there would be no opportunity for us to *apply* the lemma in the absence of a last decision point. At an intuitive level it would remain clear, for example, that no agreement can be expected after Ξ^* (Proposition 3.6), but without a deadline this conclusion would no longer follow from iterative dominance — at least absent some further modification to our model not attempted here.

Another natural question is whether the consequences of default (indicated by ω) cannot simply be subsumed into the set A. This would allow the bargainers to abandon their attempts to reach an agreement at any point during the negotiation, and thus save on waiting costs if disagreement is thought to be likely anyway (or even desirable).

One response to this question is that if we impose $\omega \in A$, then it is far from clear which coalitions should be given the right to force this alternative. If we wish to allow the agents to "walk away from the negotiating table" unilaterally, then $\forall i \in I$ we must have $\rho(\omega_i, \beta_{-i}) = \omega$. But, in view of Assumption 2.2A, this will give each agent a veto over every alternative; i.e., it is consistent only with the unanimity rule. The message here is that our basic model is not rich enough to accommodate outside options, and that these will need to be introduced explicitly by changing the extensive form.

A second response to the proposal to let $\omega \in A$ is that there are plenty of situations in which bargainers will not be able to end their discussions prematurely. For example, when negotiating authority is delegated and the delegates bear the costs of delay (as in the trade liberalization scenario of Section 1.2), early termination of the attempt to find an agreement will not typically be permitted. Similarly, a jury does not have the power to declare itself to be deadlocked; this is a decision for the presiding judge, who may well order a return to the jury room for further deliberations. And even when sovereign entities are bargaining amongst themselves in the absence of any overarching authority that can compel participation, they may not be able to irrevocably *commit* to abandoning their negotiation before the opportunity for an agreement disappears.

5.2 Static and dynamic bargaining

When introducing Terminal Solvability in Section 3.3, we referred to a bargaining problem as "static" if it has no significant intertemporal element — as opposed to "dynamic" problems that do have such an element.[16] And as already observed, our dynamic model reduces to a static problem upon the arrival of the final decision point $n\Delta = 1$. Taking this point of view one step further, we can think of the overall model as consisting of $n+1$ static problems corresponding to the decision points and linked by the passage of time. While the consequences of disagreement in the last of these problems are captured in the exogenous default outcome $\langle \omega, 1 \rangle$, the continuation outcomes of the earlier n problems will be endogenous.

The reinterpretation of dynamic bargaining as repeated static bargaining leads to a new perspective on our theory. On this reading our backward induction lemma becomes the statement that static bargaining problems in which no more than one alternative is viable are dominance solvable. Consequently, a dynamic problem will be solvable if each of its constituent static problems has this characteristic.[17] And the conditions used in Theorem 3.16 can then be seen to supply exactly what is needed to guarantee solvability by means of the fact expressed in the lemma.[18]

This intuition for why our conditions are *sufficient* for dominance solvability also

[16]In this usage static does not imply instantaneous: If we lock two bargainers in a cell with a blank contract and declare that we will return in one hour to collect it, the resulting problem will be a static one since no benefit arises from agreeing earlier rather than later during the period of confinement.

[17]This intuition is reminiscent of Moulin's [16, pp. 1342–1343] observation that a certain variety of dynamic "voting scheme" will be dominance solvable whenever each of the static schemes it contains has this property. The main difference here is that Moulin requires his static schemes to be solvable for any preferences on the part of the voters (like perfect-information game forms), whereas we consider only the particular preference profile induced by backward induction on the larger dynamic problem.

[18]Viz.: Terminal Solvability is needed to ensure that at most one alternative is viable in the final static problem (since latest acceptance points can stack up at the deadline), while Core Membership is needed to ensure that in static problems prior to decision point $k^*\Delta$ (see Section 3.3) no alternative is viable other than the unique $a^* \in A$ for which $\Xi(a^*) = \Xi^*$.

helps to clarify why they cannot be *necessary.* Lemma 3.2 says nothing about static problems in which more than one alternative is viable, and in fact these may or may not be solvable.[19] By constructing a dynamic problem out of solvable static problems, at least one of which is of this "multi-viable" type, we can therefore create a game that violates either of our conditions but nevertheless exhibits solvability.

In general, of course, multi-viable static problems will *not* be dominance solvable and so (when our conditions fail) some further assumption regarding the resolution of these problems will be required if we are to make definite predictions.[20] An example of such an assumption is temporal monopoly, which in this context would state that at each decision point some agent is given the right to make a "take it or leave it" proposal. To dispense bargaining power by fiat in this fashion may seem arbitrary and undesirable. But no alternative assumption springs to mind as an obvious improvement, and thus it appears that we may have to either lower our ambitions for bargaining theory or live with the "monopoly problem" (see Section 1.1).[21]

Similarly, and in conclusion, we can use our new perspective to better understand the "multiplicity problem" afflicting many bargaining models. While not the only factor relevant here, the presence of multi-viable static bargaining problems within a dynamic model can contribute to the proliferation of equilibria. Multiple viable alternatives set up a coordination game among the agents, so unless the extensive form employed imposes a theory of static bargaining comparable to temporal monopoly, multiple equilibria robust to refinements (a feature of coordination games generally) will be unavoidable. And since dynamic considerations such as delay costs can have no bearing on static problems, the idea that relative patience confers bargaining power — an insight central to the modern literature on negotiations — cannot help us here.

[19] For an example of the solvable case, let $I = \{1, 2, 3\}$ and $A = \{a, b\}$; endow the agents with the preferences $a \succ_1 b \succ_1 \omega$, $a \succ_2 b \succ_2 \omega$, and $b \succ_3 \omega \succ_3 a$ (where ω represents the continuation outcome); and suppose that the support of at least two agents is required to force either alternative. Here both a and b are viable with respect to ω, but the simultaneous voting game is dominance solvable: Neither 1 nor 2 will vote for β (i.e., continuation), hence 3 will vote for b, hence 1 and 2 will vote for a, and hence a will be selected.

[20] Analogously, in the context of incomplete information, Kennan and Wilson [13, pp. 50–55] stress the indeterminacy of incentive compatibility analysis (e.g., Myerson [17]) as compared to the "fairly specific predictions" obtainable by imposing a rigid structure on negotiations.

[21] Recall that low ambitions were the norm until fairly recently. Von Mises [26, p. 327], for example, argued that in bilateral barter scenarios "the ratio of exchange is determined only within broad margins. [T]heory cannot determine [exactly which] ratio will be established. All that it can assert with regard to such exchanges is that they can be effected only if each party values what he receives more highly than what he gives away." As applied to *static* bargaining problems, this view remains perfectly defensible.

References

[1] Dilip Abreu and Faruk Gul. Bargaining and reputation. *Econometrica*, 68(1):85–117, January 2000.

[2] Jeffrey S. Banks and John Duggan. A bargaining model of collective choice. *American Political Science Review*, 94(1):73–88, March 2000.

[3] David P. Baron and John A. Ferejohn. Bargaining in legislatures. *American Political Science Review*, 83(4):1181–1206, December 1989.

[4] David H. Blackwell and Meyer A. Girshick. *Theory of Games and Statistical Decisions.* Wiley, New York, 1954.

[5] Adam M. Brandenburger and H. Jerome Keisler. Epistemic conditions for iterated admissibility. Unpublished, November 2003.

[6] Kalyan Chatterjee and Larry Samuelson. Perfect equilibria in simultaneous-offers bargaining. *International Journal of Game Theory*, 19(3):237–267, 1990.

[7] Seok-ju Cho and John Duggan. Bargaining foundations of the median voter theorem. Unpublished, December 2005.

[8] Eddie Dekel. Simultaneous offers and the inefficiency of bargaining: A two-period example. *Journal of Economic Theory*, 50(2):300–308, April 1990.

[9] Christian Ewerhart. Ex-ante justifiable behavior, common knowledge, and iterated admissibility. Unpublished, August 2002.

[10] Drew Fudenberg and Jean Tirole. *Game Theory.* MIT Press, Cambridge MA, 1991.

[11] Rodney J. Gretlein. Dominance solvable voting schemes: A comment. *Econometrica*, 50(2):527–528, March 1982.

[12] Rodney J. Gretlein. Dominance elimination procedures on finite alternative games. *International Journal of Game Theory*, 12(2):107–113, June 1983.

[13] John Kennan and Robert B. Wilson. Bargaining with private information. *Journal of Economic Literature*, 31(1):45–104, March 1993.

[14] David M. Kreps. *A Course in Microeconomic Theory.* Princeton University Press, Princeton NJ, 1990.

[15] Leslie M. Marx and Jeroen M. Swinkels. Order independence for iterated weak dominance. *Games and Economic Behavior*, 18(2):219–245, February 1997.

[16] Hervé Moulin. Dominance solvable voting schemes. *Econometrica*, 47(6):1337–1351, November 1979.

[17] Roger B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47(1):61–74, January 1979.

[18] John F. Nash, Jr. Two-person cooperative games. *Econometrica*, 21(1):128–140, January 1953.

[19] Motty Perry and Philip J. Reny. A non-cooperative bargaining model with strategically timed offers. *Journal of Economic Theory*, 59(1):50–77, February 1993.

[20] Ariel Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–110, January 1982.

[21] Jozsef Sakovics. Delay in bargaining games with complete information. *Journal of Economic Theory*, 59(1):78–95, February 1993.

[22] Alp Simsek and Muhamet Yildiz. Durable bargaining power and stochastic deadlines. Unpublished, January 2008.

[23] Lones Smith and Ennio Stacchetti. Aspirational bargaining. Unpublished, November 2003.

[24] Ingolf Stahl. *Bargaining Theory*. Economic Research Institute, Stockholm, 1972.

[25] John Sutton. Non-cooperative bargaining theory: An introduction. *Review of Economic Studies*, 53(5):709–724, October 1986.

[26] Ludwig von Mises. *Human Action*. Contemporary Books, Chicago, 1966.